Genomics 102 (2013) 157–162

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Multiplexed deep sequencing analysis of *ALK* kinase domain identifies resistance mutations in relapsed patients following crizotinib treatment

Donghui Huang [a,*,1], Dong-Wan Kim [b,1], Athanasios Kotsakis [c], Shibing Deng [a], Paul Lira [a], Steffan N. Ho [a], Nathan V. Lee [a], Pamela Vizcarra [a], Joan Q. Cao [a], James G. Christensen [a], Tae Min Kim [b], Jong-mu Sun [d], Jin Seok Ahn [d], Myung-Ju Ahn [d], Keunchil Park [d,**], Mao Mao [a,*]

[a] *Pfizer Oncology, San Diego, CA, USA*
[b] *Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea*
[c] *Oncology Department, University Hospital of Heraklion, Crete, Greece*
[d] *Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea*

ARTICLE INFO

Article history:
Received 6 September 2012
Accepted 11 February 2013
Available online 20 February 2013

Keywords:
Crizotinib
Resistance
Mutation
Deep sequencing

ABSTRACT

The recently approved *ALK* kinase inhibitor crizotinib has demonstrated successful treatment of metastatic and late stage *ALK* fusion positive non-small cell lung cancer (NSCLC). However, the median duration of clinical benefit is ~10–11 months due to the emergence of multiple and simultaneous resistance mechanisms in these tumors. Mutations in the *ALK* kinase domain confer resistance to crizotinib in about one-third of these patients. We developed a multiplex deep sequencing method using semiconductor sequencing technology to quickly detect resistance mutations within the *ALK* kinase domain from tumor biopsies. By applying a base-pair specific error-weighted mutation calling algorithm (BASCA) that we developed for this assay, genomic DNA analysis from thirteen relapsed patients revealed three known crizotinib resistance mutations, C1156Y, L1196M and G1269A. Our assay demonstrates robust and sensitive detection of *ALK* kinase mutations in NSCLC tumor samples and aids in the elucidation of resistance mechanisms pertinent to the clinical setting.

1. Introduction

Crizotinib was approved by US FDA in 2011 as a first-in-class, oral, small-molecule anaplastic lymphoma kinase (*ALK*) inhibitor for the treatment of *ALK* fusion positive advanced non-small cell lung cancer (NSCLC). Phase I and II clinical trials of crizotinib showed an objective response rate of 50–61% with median response durations of 42–48 weeks, and a median progression-free survival (PFS) of 10–11 months [1,2]. The frequency of *ALK* rearrangements in the overall population of NSCLC patients is 3–5% or ~45,000 patients worldwide each year [3]. The rearranged *ALK* gene and its fusion partner mediate ligand-independent oligomerization of *ALK* and constitutive kinase activation. Tumors possessing *ALK* rearrangement such as *EML4-ALK* become oncogenically addicted to *ALK* signaling, and at the same time make them vulnerable to *ALK* inhibitors [4–6]. However, many patients with initial encouraging responses later developed resistance. About one-third of the patients with acquired resistance to crizotinib have secondary mutations in the *ALK* tyrosine kinase domain or amplification of the *ALK* fusion gene [7–9]. *ALK* resistance mutations vary widely within the tyrosine kinase domain. So far, nine different acquired resistance mutations have been identified from crizotinib treated NSCLC patients, including 1151Tins, L1152R, C1156Y, F1174C, L1196M, G1202R, D1203N, S1206Y and G1296A [7–10,17]. Another third of the crizotinib-resistant patients' tumors possess bypass mechanisms via activation of EGFR, KIT or KRAS, and the last third of refractory patients have undetermined resistance mechanisms [8,9]. Efforts are underway to develop second generation *ALK* inhibitors that are more potent towards *ALK* kinase domain mutations. Therefore, assays that can detect *ALK* kinase mutations from tumor biopsies may improve clinical outcomes. Several highly sensitive platforms exist to detect low frequency mutations in very heterogeneous biopsy samples. Of these technologies including primer extension [8,11], allele specific PCR [12], pyrosequencing [13], BEAMing [14] and droplet digital PCR [15], next generation deep sequencing technology stands out as the optimal interrogation platform because it can detect novel mutations and requires a small amount of specimens. We utilized the Ion Torrent Personal Genome Machine (PGM) to deep sequence the entire *ALK* kinase domain of tumor DNA from NSCLC patients who have relapsed after crizotinib treatment. Our assay identified three previously reported resistance mutations in three of thirteen patients' tumor biopsy samples.

* Corresponding authors at: Pfizer Oncology, 10724 Science Center Drive, San Diego, CA 92121.
** Correspondence to: K. Park, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 135-710, Seoul, Korea.
E-mail addresses: donghuistephen.huang@pfizer.com (D. Huang), kpark@skku.edu (K. Park), mao.mao@pfizer.com (M. Mao).
[1] D. Huang and D.W. Kim contributed equally to this study.

0888-7543/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ygeno.2013.02.006

Table 1
Detection limit of mutation G1269A with the presence of wild type *ALK*.

% G1269A allele	Total reads	Number of reads with G1269A allele
20.0%	2731	546 (20.0%)
10.0%	2291	180 (7.9%)
5.0%	2138	100 (4.7%)
2.5%	2769	82 (2.9%)
1.25%	2500	46 (1.8%)
0%	2322	0 (0%)

2. Results

The assay was used to detect *ALK* kinase domain mutations from crizotinib-resistant NCI-H3122 clones derived in vitro. Using this technology, a G1269A mutation was identified in clone 1-07 with 20% allele frequency and a total of 2731 reads of bidirectional coverage. This mutation was also confirmed by Sanger sequencing (data not shown). With an abundance of gDNA from a resistant cell line, we proceeded to determine the sensitivity of this assay running on an Ion Torrent deep sequencing platform using spike-in titration experiments. Genomic DNA from the crizotinib-resistant clone 1-07 was spiked into different amounts of a human male normal genomic DNA. The amount of G1269A mutation ranged from 1.25%–20% of the total *ALK* genes presence. The full sensitivity range is summarized in Table 1 where the G1269A mutation was detectable at the level as low as 1.25%. However, average sequencing noise level ~0.13% was determined and observed by examining ~67 base-pairs flanking region of the codon G1269 in exon 25 from the human male normal sample sequencing reads with a standard deviation of 0.21%. Therefore, we should be able to call mutation G1269A above 0.8% (average noise plus 3× standard deviation) threshold with 99% confidence. Overall, the highest noise level in all exons 21–27 observed in the human male normal sample was about 3%.

Each patient DNA sample was amplified with barcoded fusion primer pairs and bi-directionally sequenced on two Ion Torrent 314 chip in two independent sequencing runs (Fig. 1). Since the sequencing adapters and barcodes were built into the target specific primers, sequencing libraries were formed during the PCR step (Fig. 1). PCR amplicons were pooled into a single tube for deep sequencing on an Ion Torrent PGM. Results showed an excess of 500× coverage from either direction covering exons 21–27 of *ALK* kinase domain in all samples. Mean quality score Q20 read length was approximately 148 bps, and mutations were identified with bases different from references but of whose quality scores were higher and equal to 20. Thirteen tumor samples were collected from the patients who were relapsed post crizotinib treatment. Two of the patients included archived pre-treatment FFPE specimens (Patients 5 and 6). Source materials from these patients included a lung primary tumor resection, metastatic liver tumor biopsy, metastatic bone tumor excision, metastatic adrenal tumor excision, bronchoscopic biopsy and pleural fluid (Table 2). Each sample was sectioned, H&E stained, and examined for tumor cell content. Pathology reports showed tumor content ranging from 10–90%. All *ALK* kinase domain exons from all thirteen patients including archived FFPE materials were amplified successfully. Each sample was amplified with barcoded fusion primer pairs and bi-directionally sequenced on an Ion Torrent PGM (Fig. 1).

Instead of using an arbitrary cutoff across the seven exons of *ALK* kinase domain sequenced by PGM (~1 kb), we developed a base-pair specific error-weighted mutation calling algorithm (BASCA) for detecting mutations by assessing the noise level at each single base-pair in order to increase the accuracy of mutation calling. Three unique mutations were identified from three tumors by BASCA after filtered out known SNPs and manual inspection to eliminate two false positive calls in the sequencing error-prone regions. Patient 2's bone metastatic tumor sample revealed a known mutation G1269A with 15% allele frequency (Fig. 2A). This call was supported by a total of 10,590 reads with bidirectional coverage. Interestingly, Patient 2 partially responded to crizotinib treatment but relapsed after 12 months with metastatic tumors found in the bone. For Patient 7, normal and tumor regions were dissected from a single paraffin block of a liver biopsy. No mutations were detected in the normal liver tissue. However, a mutation, C1156Y, was detected in the metastatic tumor sample with 14% allele frequency. This finding was supported with 1973 reads including bidirectional coverage (Fig. 2B). Patient 7 was an initial crizotinib responder who relapsed after 8 months with liver metastasis. In Patient 8's lung core biopsy tumor sample, the "gatekeeper" mutation L1196M was identified with 12% allele frequency (Fig. 2C). This call was supported by a total of 12,269 reads with bidirectional coverage. Patient 8 achieved partial response to crizotinib for 7 months as the 7th line of therapy but progressed in the liver and brain.

No mutations were detected in the *ALK* kinase domain in the rest ten patients. In addition, no preexisting mutations were found in the archived materials from Patients 5 and 6 prior to crizotinib treatment (Table 2). Tumor samples from Patients 1, 2, 5 and 6 were also examined by Sanger sequencing and allele specific PCR [18]. These results were concordant with the Ion Torrent PGM deep sequencing data, i.e. G1269A mutation was detected in Patient 2's pleural effusion sample (a different material from the patient) and no mutation was detected in the samples from Patients 1, 5 and 6 [18].

3. Discussion

After the recent approval of crizotinib by US FDA, most *ALK* fusion positive NSCLC patients will have significant clinical benefit from the drug. However, it is plausible that acquired resistance may ensue like other targeted tyrosine kinase inhibitor (TKI) therapeutics [16]. In this study, we successfully developed a molecular assay capable of detecting three known resistant mutations in the *ALK* kinase domain

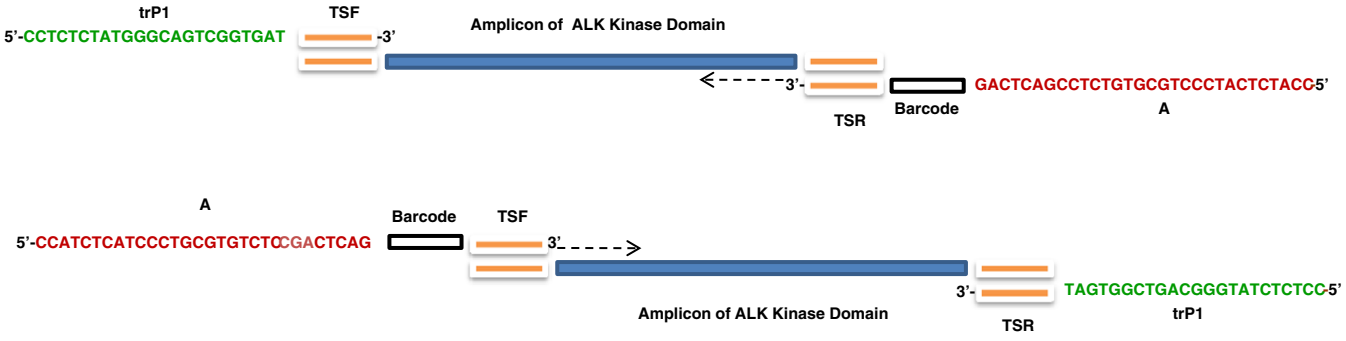

Fig. 1. Barcoded fusion primers. One fusion primer pair has the A adapter (A) region followed by a barcode sequence, then followed by the template-specific forward primer (TSF) in the proximal end of the targeted *ALK* exon sequence, and the other has the truncated P1 adapter (trP1) region followed by the template-specific reversed primer (TSR) in the distal end of the target *ALK* exon sequence. The other fusion primer pair has the adapter sequences A and trP1 swapped. Two primer pairs per exon region generate two libraries to enable bidirectional sequencing of each exon region. This is the same for exons 21–27 of *ALK* kinase domain.

Table 2
Summary of sequencing results of all materials from thirteen patients.

Patient ID	Sampling	Sample type	Tumor content (%)	Mutation	Confirmed by
1	Post treatment	Liver metastasis	50	Not detected	Sanger*
2	Post treatment	Bone metastasis	90	G1269A (15%)	Sanger and AS-PCR*
3	Post treatment	Pleural fluid	70	Not detected	
4	Post treatment	Adrenal metastasis	50	Not detected	
5	Pretreatment	Lung primary	40	Not detected	Sanger*
	Post treatment	Bronchoscopic biopsy	30	Not detected	
6	Pretreatment	Bone metastasis	70	Not detected	Sanger*
	Post treatment	Liver metastasis	80	Not detected	
7	Post treatment	Normal liver tissue	0	Not detected	
	Post treatment	Liver metastasis	70	C1156Y (14%)	
8	Post treatment	Liver metastasis	70	L1196M (12%)	
9	Post treatment	Lung core biopsy	N/A	Not detected	
10	Post treatment	Chest wall soft tissue	N/A	Not detected	
11	Post treatment	Lung core biopsy	50	Not detected	
12	Post treatment	Neck lymph node	80	Not detected	
13	Post treatment	Adrenal metastasis	10	Not detected	

*confirmation performed by Sanger sequencing and allele specific PCR (AS-PCR) [18].

in three out of thirteen (23.1%) relapsed patients. This is also in agreement with the other studies that about one-third of the crizotinib acquired resistant tumors harbor secondary mutations. These three mutations, C1156Y, L1196M and G1269A, were previously reported to be present in crizotinib-refractory patients [7,8]. C1156Y was first reported alongside the gatekeeper mutation L1196M, in a Japanese patient who relapsed after 5 months of crizotinib therapy. C1156Y is positioned adjacent to the N-terminal helix αC and induces conformational changes that negatively affect the binding activity of crizotinib to *ALK* [7]. G1269A was previously found in two patients who confer resistance to crizotinib. Both patients had partial responses (PR) after crizotinib treatment, and later developed acquired resistance [8]. The G1269A residue is positioned at the end of the narrow ATP-binding pocket of *ALK* where alanine significantly compromises crizotinib binding due to steric hindrance. The G1269A finding in Patient 2's bone metastasis and pleural effusions was the third case of a relapsed patient reported to date with this particular mutation. Given these finding, it is conceivable that ATP-binding pocket associated transformations are a more frequent mechanism of crizotinib resistance in relapsed patients. Wide spectrum of resistant mutations in the *ALK* kinase domain is one of the major resistance mechanisms. Other possible resistance mechanisms include the activation of compensatory signaling pathways that are unaffected by *ALK* inhibitors. Therefore, once a patient develops acquired resistance to a particular therapy, it is essential to quickly assess the patient's tumor cell profile to tailor specific treatment options. Therein lies the need for an accurate and sensitive test since one third of relapsed patients will have secondary mutations in their *ALK* gene. In a recent study, new *ALK* inhibitors under active clinical development have differential potencies against the different resistance mutations [9]. With the knowledge of a patient's *ALK*-resistant mechanism, we can strategically treat the patients with the second generation *ALK* inhibitors or combinations with hsp90 inhibitors, chemotherapies, or other TKIs to overcome acquired resistance and improve clinical outcome.

We have developed a sensitive assay that can detect *ALK* secondary mutations in FFPE samples from various tissue origins. A variety of clinical samples collected from relapsed patients will be the source for future mutation analysis. However, several challenges remain.

Table 3
Barcoded fusion primer sequences.

Primer name	Sequence
A-Barcode-Exon27F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – GG TGT GTC TAT ATC CAT CTC CA
A-Barcode-Exon27R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – AG CAT ATG TGG CTC TGG ATA
A-Barcode-Exon26F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – GC AGG GCA GAT GCT TAA T
A-Barcode-Exon26R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – GG AGG ATG ATG GCT GAC TT
A-Barcode-Exon25F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – GC ATT TCC TTT CTT CCC AG
A-Barcode-Exon25R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – CA CAC CCC ATT CTT GAG G
A-Barcode-Exon24F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – CA CTG ACA AGC TCC TCG TC
A-Barcode-Exon24R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – AG CGA CAG GAT GAC AGG A
A-Barcode-Exon23F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – GA CTC AGC TCA GTT AAT TTT GG
A-Barcode-Exon23R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – CA GCA AAG ACT GGT TCT CAC
A-Barcode-Exon22F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – AG TTC TCA GCT CAC AGC CT
A-Barcode-Exon22R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – AC CCT CTC CAG GTT CTT TGG
A-Barcode-Exon21F	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – CC TCA TTA TTG TGG CCT GT
A-Barcode-Exon21R	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG – BARCODE – AA GGG CAG GCT CAA GAG T
trP1-Exon27F	CCT CTC TAT GGG CAG TCG GTG ATG GTG TGT CTA TAT CCA TCT CCA
trP1-Exon27R	CCT CTC TAT GGG CAG TCG GTG ATA GCA TAT GTG GCT CTG GAT A
trP1-Exon26F	CCT CTC TAT GGG CAG TCG GTG ATG CAG GGC AGA TGC TTA AT
trP1-Exon26R	CCT CTC TAT GGG CAG TCG GTG ATC CCA GGA GCA CCA CCT TAT G
trP1-Exon25F	CCT CTC TAT GGG CAG TCG GTG ATG CAT TTC CTT TCT TCC CAG
trP1-Exon25R	CCT CTC TAT GGG CAG TCG GTG ATC ACA CCC CAT TCT TGA GG
trP1-Exon24F	CCT CTC TAT GGG CAG TCG GTG ATC ACT GAC AAG CTC CTC GTC
trP1-Exon24R	CCT CTC TAT GGG CAG TCG GTG ATA GCG ACA GGA TGA CAG GA
trP1-Exon23F	CCT CTC TAT GGG CAG TCG GTG ATG ACT CAG CTC AGT TAA TTT TGG
trP1-Exon23R	CCT CTC TAT GGG CAG TCG GTG ATC AGC AAA GAC TGG TTC TCA C
trP1-Exon22F	CCT CTC TAT GGG CAG TCG GTG ATA GTT CTC AGC TCA CAG CCT
trP1-Exon22R	CCT CTC TAT GGG CAG TCG GTG ATA CCC TCT CCA GGT TCT TTG G
trP1-Exon21F	CCT CTC TAT GGG CAG TCG GTG ATC CTC ATT ATT GTG GCC TGT
trP1-Exon21R	CCT CTC TAT GGG CAG TCG GTG ATA AGG GCA GGC TCA AGA GT

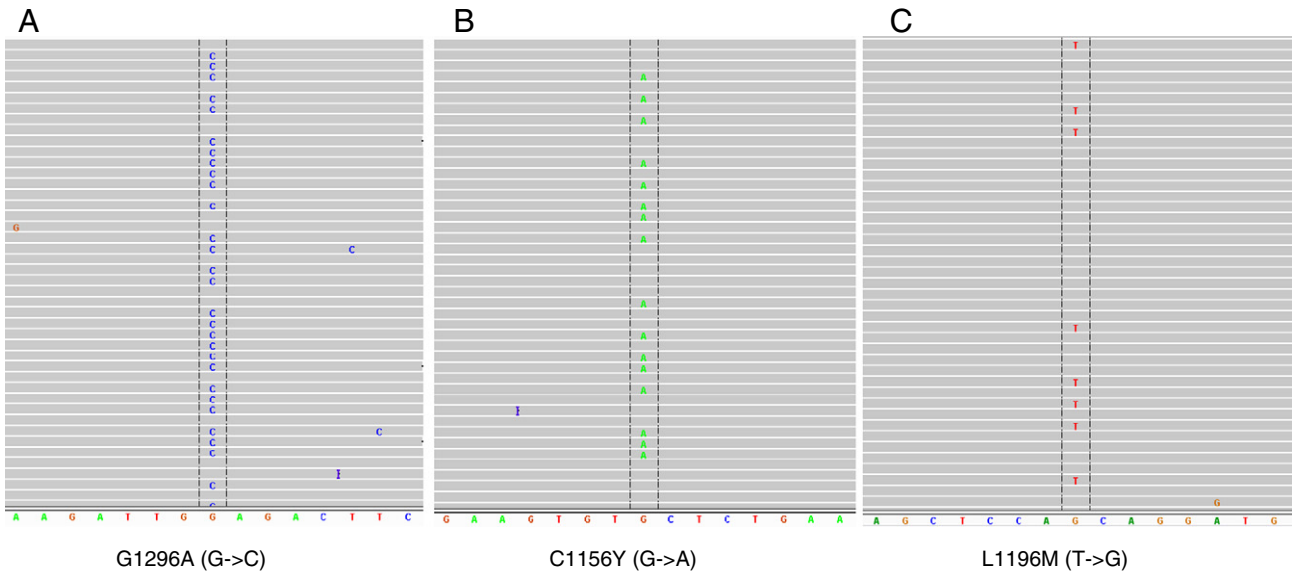

Fig. 2. Snapshot views of IVG viewer of mutations L1196M, C1156Y and G1296A. Panel A: mutation G1296A from Patient 2's tumor sample. Panel B: mutation C1156Y from Patient 7's dissected tumor sample. Panel C: mutation L1196M from Patient 8's tumor sample.

Rebiopsying tumors in relapsed patients may be difficult because they are usually in critical condition. Secondly, tumor content or the ratio of mutant and normal cells could affect the detection limit by diluting the mutant allele. Thirdly, the DNA in FFPE samples could be poorly preserved, degraded and chemically-crosslinked. We have shown that it is possible to extract genomic DNA from FFPE samples and amplify it for downstream sequencing. DNA from large resected tumors was isolated with a column based DNA extraction method. Small tissue sections or needle biopsies were extracted with a commercial proteinase K lysis buffer kit. In this study, we successfully amplified all amplicons using extracted genomic DNA from various sample types including pleural fluid and bronchoscopic biopsy in FFPE. We were able to detect three mutations, C1156Y, L1196M and G1269A, with an allele frequency between 12% and 15% in three patient tumor samples. The assay was sensitive to detect mutant allele at 1.25% level when we used a known mutation G1269A for a series of titration experiments.

We successfully developed a simple strategy to assess the mutation status of crizotinib-resistant tumor. Eight sets of PCR primers containing barcodes and adapter sequences were placed into a 96 well plate for PCR. All amplicons were pooled and directly added into Ion Torrent sequencing workflow. The entire process using 8 to 16 samples per sequencing run can be completed in three working days. We demonstrated a sensitive and accurate deep sequencing approach that can identify secondary *ALK* mutation from various forms of clinical samples. This assay could be used in the clinical setting to help physicians quickly choose treatment options to overcome crizotinib resistance and improve patient outcome. Our base-pair specific error-weighted mutation calling algorithm (BASCA) could also be applied to the other targeted deep sequencing assays.

4. Materials and methods

4.1. Patients

Thirteen *ALK* fusion positive NSCLC patients were investigated after acquired resistance was developed during the course of crizotinib treatment. Tumor samples were collected prior to the treatment and/or after relapse and tested using approved protocols concordant with Institutional Review Board in the three hospitals.

4.2. Histologic examination of FFPE tumor samples

A 5 μm section was cut from the provided FFPE block and placed onto a positive charged slide to air dry for 30 min and then heat fixed at 60 °C for 60 min. The slide was cooled, deparaffinized, and stained with hematoxylin and eosin using the Leica Select Tech H&E Staining System (Leica Microsystems Inc., Buffalo Grove, IL) and the Leica Autostainer XL (Leica Microsystems Inc., Buffalo Grove, IL). Tumor content was estimated during histological examination.

4.3. Cell line and normal genomic DNA for titration experiments

NCI-H3122 cells (ATCC, Manassas, VA) were cultured in Gibco RPMI 1640 (Cat# 22400-071, Life Technologies, Grand Island, NY), 10% FBS (Cat# 10437-028, Life Technologies, Grand Island, NY), and, 1% Penicillin Streptomycin (Cat# 15140-148, Life Technologies, Grand Island, NY) under selection of increasing concentrations of crizotinib increase of 2 fold every 2–3 weeks, starting at 100 nM and ending at 1 μM after 10 weeks. These cells were then clonally selected with cloning disks and indefinitely maintained at 1 μM of crizotinib. Of the resistance clones we isolated, clone 1-07 was identified to possess the G1269A mutation within the *EML4-ALK* kinase domain. Genomic DNA was extracted using a Qiagen DNeasy Blood & Tissue Kit (Cat# 69504, Qiagen Inc., Valencia, CA). Normal human male gDNA was obtained from Life Technologies (Cat# 360486, Life Technologies, Grand Island, NY).

Table 4
Barcode sequences.

Barcode	Sequence
1	ACGAGTGCGT
2	ACGCTCGACA
3	AGACGCACTC
4	TCGATAGTGA
5	ATCAGACACG
6	ATATCGCGAG
7	CGTGTCTCTA
8	CTCGCGTGTC

4.4. Titration experiment to determine assay sensitivity

Clone 1-07 genomic DNA was diluted and titrated down to varying percentages at 100%, 50%, 25%, 12.5%, 6.25% and 0% relative to the normal human male gDNA. All samples were PCR-amplified and barcoded with barcodes 1–6 and sequenced on an Ion Torrent PGM.

4.5. Genomic DNA extraction for ALK exon PCR

Genomic DNA was isolated from FFPE tumor sections with a Qiagen QIAamp® DNA FFPE Tissue Kit (Cat# 56404, Qiagen Inc., Valencia, CA) according to the manufacturer's protocol. Alternatively, DNAs from needle biopsies and small dissected tumor sections were extracted with the Arcturus® PicoPure® DNA Extraction Kit (Cat# KIT0103, Life Technologies, Grand Island, NY) using the manufacturer's "DNA Extraction Protocol: CapSure HS LCM Cap Samples". Exons 21 to 27 corresponding to the *ALK* tyrosine kinase domain were PCR-amplified and barcoded from 20 ng QIAamp isolated genomic DNA or 10 μl PicoPure® DNA Extraction lysates with the Roche FastStart High Fidelity PCR system (Cat# 03553361001, Roche Applied Science, Indianapolis, IN). PCR conditions were set to the manufacturer's protocol. PCR primers and barcode sequences are listed in Table 3 and 4. A primer master plate was created to facilitate multichannel pipetting for PCR setup. A no template control was included in the PCR setup using barcode 1 primer sets.

4.6. Sequencing on Ion Torrent Personal Genome Machine™ (PGM)

The DNA concentration for each amplicon was quantified on an Agilent Bioanalyzer using the DNA 1000 chip (Cat# 5067-1504, Agilent Technologies, Inc., Santa Clara, CA) according to the manufacturer's protocol. The barcoded PCR amplicon templates were equimolarly pooled and sequenced bi-directionally on an Ion Torrent PGM (Life Technologies, Grand Island, NY) according to manufacturer's protocols. A total of 2×10^8 DNA molecules library was constructed in each sequencing run. The barcode-specific sample data was analyzed by Life Technologies Torrent Suite version 2.2. Exported BAM files were reviewed using Broad Institute's Integrative Genomics Viewer (IGV) version 2.1. In IGV, the "Coverage allele-freq threshold" was set to 0.01, the "Shade mismatched bases by quality" was set to 0 to 20.

4.7. BASCA algorithm and data analysis

Aligned sequence data were processed using Samtools [19]. The results were piled up to obtain coverage data for both forward and reverse strands as well as reads for individual nucleotides at each base location. Mutation rate was defined as the sum of reads from non-reference nucleotides divided by the total coverage at each base location, expressed in percent. A mutation is detected at base location i if its lower 95% confidence interval bound is above a threshold that is defined as

$$M_{ti} = \max(M_{ci}, M_b)$$

where M_{ci} is a base-specific threshold, based on the distribution of mutation rate at base location i across samples, $M_{ci} = (\hat{x}_i + 4s_i)^2$, where $\hat{x}_i$ is the median of the square root transformed mutation rate across all samples, and s_i is the robust standard deviation of the square root transformed mutation rate, which is estimated using median absolute deviation from median (MAD). Square root transformation helped to normalize the data of small proportions. The use of 4 standard deviations above median in the threshold determination helped to prevent potential false positives as mutation is a rate event. M_b is a global background threshold for the mutation rate across all bases and samples. It provides a protection of a small base-specific threshold due to poor estimate of variability from relatively small sample size. It was defined as the upper 95% confidence bound (97.5 percentile) of the mutation rate in this data. The threshold M_{ci} provides a measure of noise level of the mutation rate at base i, and we assumed that most of the data was not mutated. For quality purpose, data with coverage below 200x were excluded. Mutation data were not of our interest if there was an extreme strand bias (one strand has coverage of more than 10 times higher than the other), inconsistent mutation between two strands, or at a base where a common SNP was reported. All data were analyzed in R v2.13.1 9 (r-project.org).

Acknowledgments

Technical assistance from Dr. Keith Ching is gratefully acknowledged.

References

[1] D.R. Camidge, Y. Bang, E.L. Kwak, A.T. Shaw, A.J. Iafrate, R.G. Maki, B.J. Solomon, S.I. Ou, R. Salgia, K.D. Wilner, D.B. Costa, G. Shapiro, P. LoRusso, P. Stephenson, Y. Tang, K. Ruffner, J.W. Clark, Progression-free survival (PFS) from a phase I study of crizotinib (PF-02341066) in patients with ALK-positive non-small cell lung cancer (NSCLC), J. Clin. Oncol. 29 (2011) 2501, (ASCO meeting abstracts).
[2] L. Crinò, D. Kim, G.J. Riely, P.A. Janne, F.H. Blackhall, D.R. Camidge, V. Hirsh, T. Mok, B.J. Solomon, K. Park, S.M. Gadgeel, R. Martins, J. Han, T.M. De Pas, A. Bottomley, A. Polli, J. Petersen, V.R. Tassell, A.T. Shaw, Initial phase II results with crizotinib in advanced ALK-positive non-small cell lung cancer (NSCLC): profile 1005, J Clin Oncol. 29 (2011) 7514, (Suppl., ASCO meeting abstracts).
[3] R.H. Palmer, E. Vernersson, C. Grabbe, B. Hallberg, Anaplastic lymphoma kinase: signalling in development and disease, Biochem. J. 420 (2009) 345–361.
[4] M. Soda, S. Takada, K. Takeuchi, Y.L. Choi, M. Enomoto, T. Ueno, H. Haruta, T. Hamada, Y. Yamashita, Y. Ishikawa, Y. Sugiyama, H. Mano, A mouse model for EML4-ALK-positive lung cancer, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 19893–19897.
[5] J.P. Koivunen, C. Mermel, K. Zejnullahu, C. Murphy, E. Lifshits, A.J. Holmes, H.G. Choi, J. Kim, D. Chiang, R. Thomas, J. Lee, W.G. Richards, D.J. Sugarbaker, C. Ducko, N. Lindeman, J.P. Marcoux, J.A. Engelman, N.S. Gray, C. Lee, M. Meyerson, P.A. Jänne, EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer, Clin. Cancer Res. 14 (2008) 4275–4283.
[6] U. McDermott, A.J. Iafrate, N.S. Gray, T. Shioda, M. Classon, S. Maheswaran, W. Zhou, H.G. Choi, S.L. Smith, L. Dowell, L.E. Ulkus, G. Kuhlmann, P. Greninger, J.G. Christensen, D.A. Haber, J. Settleman, Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors, Cancer Res. 68 (2008) 3389–3395.
[7] Y.L. Choi, M. Soda, Y. Yamashita, T. Ueno, J. Takashima, T. Nakajima, Y. Yatabe, K. Takeuchi, T. Hamada, H. Haruta, Y. Ishikawa, H. Kimura, T. Mitsudomi, Y. Tanio, H. Mano, EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors, N. Engl. J. Med. 363 (2010) 1734–1739.
[8] R.C. Doebele, A.B. Pilling, D.L. Aisner, T.G. Kutateladze, A.T. Le, A.J. Weickhardt, K.L. Kondo, D.J. Linderman, L.E. Heasley, W.A. Franklin, M. Varella-Garcia, D.R. Camidge, Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer, Clin. Cancer Res. 18 (5) (2012) 1472–1482, (Mar 1).
[9] R. Katayama, A.T. Shaw, T.M. Khan, M. Mino-Kenudson, B.J. Solomon, B. Halmos, N.A. Jessop, J.C. Wain, A.T. Yeo, C. Benes, L. Drew, J.C. Saeh, K. Crosby, L.V. Sequist, A. John Iafrate, J.A. Engelman, Mechanisms of acquired crizotinib resistance in ALK-rearranged lung cancers, Sci Transl Med 4 (2012) 120ra17.
[10] T. Sasaki, J. Koivunen, A. Ogino, M. Yanagita, S. Nikiforow, W. Zheng, C. Lathan, J.P. Marcoux, J. Du, K. Okuda, M. Capelletti, T. Shimamura, D. Ercan, M. Stumpfova, Y. Xiao, S. Weremowicz, M. Butaney, S. Heon, K. Wilner, J.G. Christensen, M.J. Eck, K.K. Wong, N. Lindeman, N.S. Gray, S.J. Rodig, P.A. Jänne, A novel ALK secondary mutation and EGFR signaling cause resistance to ALK kinase inhibitors, Cancer Res. 71 (18) (2011) 6051–6060, (Sep 15).
[11] R.K. Thomas, A.C. Baker, R.M. Debiasi, W. Inckler, T. Laframboise, W.M. Lin, M. Wang, W. Feng, T. Zander, L. MacConaill, J.C. Lee, R. Nicoletti, C. Hatton, M. Goyette, L. Girard, K. Majmudar, L. Ziaugra, K.K. Wong, S. Gabriel, R. Beroukhim, M. Peyton, J. Barretina, A. Dutt, C. Emery, H. Greulich, K. Shah, H. Sasaki, A. Gazdar, J. Minna, S.A. Armstrong, I.K. Mellinghoff, F.S. Hodi, G. Dranoff, P.S. Mischel, T.F. Cloughesy, S.F. Nelson, L.M. Liau, K. Mertz, M.A. Rubin, H. Moch, M. Loda, W. Catalona, J. Fletcher, S. Signoretti, F. Kaye, K.C. Anderson, G.D. Demetri, R. Dummer, S. Wagner, M. Herlyn, W.R. Sellers, M. Meyerson, L.A. Garraway, High-throughput oncogene mutation profiling in human cancer, Nat. Genet. 39 (3) (2007) 347–351, (Mar).
[12] N. Thelwell, S. Millington, A. Solinas, J. Booth, T. Brown, Mode of action and application of scorpion primers to mutation detection, Nucleic Acids Res. 28 (19) (2000) 3752–3761.
[13] S. Dufort, M.J. Richard, F. de Fraipont, Pyrosequencing method to detect KRAS mutation in formalin-fixed and paraffin-embedded tumor tissues, Anal. Biochem. 391 (2) (2009) 166–168, (Aug 15).
[14] D. Dressman, H. Yan, G. Traverso, K.W. Kinzler, B. Vogelstein, Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations, Proc. Natl. Acad. Sci. U. S. A. 100 (15) (2003) 8817–8822.
[15] B.J. Hindson, K.D. Ness, D.A. Masquelier, P. Belgrader, N.J. Heredia, A.J. Makarewicz, I.J. Bright, M.Y. Lucero, A.L. Hiddessen, T.C. Legler, T.K. Kitano, M.R. Hodel, J.F. Petersen, P.W. Wyatt, E.R. Steenblock, P.H. Shah, L.J. Bousse, C.B. Troup, J.C. Mellen, D.K. Wittmann, N.G. Erndt, T.H. Cauley, R.T. Koehler, A.P. So, S. Dube, K.A. Rose, L. Montesclaros, S. Wang, D.P. Stumbo, S.P. Hodges, S. Romine, F.P.

Milanovich, H.E. White, J.F. Regan, G.A. Karlin-Neumann, C.M. Hindson, S. Saxonov, W. Colston, High-throughput droplet digital PCR system for absolute quantitation of DNA copy number, Anal. Chem. 83 (2011) 8604–8610.
[16] E. Ardini, A. Galvani, ALK inhibitors, a pharmaceutical perspective, Front Oncol. Cancer Mol. Targets Ther. 2 (17) (2012) 1–8.
[17] D.R. Camidge, R.C. Doebele, Treating ALK-positive lung cancer—early successes and future challenges, Nat. Rev. Clin. Oncol. 9 (5) (2012) 268–277.
[18] S. Kim, T.M. Kim, D.W. Kim, H. Go, B. Keam, S.H. Lee, J.L. Ku, D.H. Chung, D.S. Heo, Heterogeneity of genetic changes associated with acquired crizotinib resistance in ALK-rearranged lung cancer, J. Thorac. Oncol. (2013), (Jan 22).
[19] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000. Genome Project Data Processing Subgroup, The Sequence alignment/map (SAM) format and SAMtools, Bioinformatics 25 (2009) 2078–2079.